

# A alternância entre “mas” e “só que” nas variedades de português brasileiro de Nova Iguaçu e do Rio de Janeiro

---

DOI: <http://dx.doi.org/10.21165/el.v50i2.2960>

**Rogério Santos Júnior<sup>1</sup>**  
**Juliana Barbosa de Segadas Vianna<sup>2</sup>**

## Resumo

Investiga-se a variação entre a conjunção *mas* e a perífrase conjuncional *só que*, com a finalidade de verificar qual é o estatuto desse fenômeno variável; e delimitar as motivações que favorecem a escolha de *só que* frente à de *mas*. Adotando o referencial teórico da Sociolinguística variacionista, utilizam-se amostras do Corpus Concordância (UFRJ). De acordo com a análise dos dados, foi possível propor as seguintes ilações: (a) que a variação entre as formas está, ainda, em um estágio inicial, sendo, portanto, estável; e (b) que o uso de *só que* está condicionado, principalmente, à função de operador discursivo, ou à função de marcador de turno conversacional; e a verbos que denotam processos materiais, processos verbais e processos mentais.

**Palavras-chave:** *só que*; *mas*; variação linguística; Rio de Janeiro; Nova Iguaçu.

---

1 Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro, Brasil; [rgsantosjunior@gmail.com](mailto:rgsantosjunior@gmail.com); <https://orcid.org/0000-0002-5042-7278>

2 Universidade Federal Rural do Rio de Janeiro (UFRRJ), Nova Iguaçu, Rio de Janeiro, Brasil; [julianasegadas@gmail.com](mailto:julianasegadas@gmail.com); <https://orcid.org/0000-0002-0146-8719>

# The variation between *mas* and *só que* in the Brazilian Portuguese varieties of Nova Iguaçu and Rio de Janeiro

## Abstract

The variation between the conjunction *mas* and the conjunctival periphrasis *só que* is investigated to delimit the motivations that favor the application of *só que* in front of *mas* and to verify the status of this variable phenomenon. Adopting the theoretical reference of the Labovian sociolinguistics, the *corpora* provided by the Concordance Project, of the Federal University of Rio de Janeiro. It has been found that (a) the variation between the forms is still at an early stage and is therefore stable; and (b) the use of *só que* is conditioned, mainly, to the following contexts: the function of a discursive operator or a conversational shift marker; to verbs that denote material processes, verbal processes and mental processes.

**Keywords:** *só que*; *mas*; linguistic variation; Rio de Janeiro; Nova Iguaçu.

## Introdução

Neste trabalho, são investigadas duas estratégias, variantes entre si, utilizadas pelos falantes para sequenciar segmentos (oracionais ou não); para criar cadeias subtópicas no discurso (função de operador discursivo); e para demarcar as mudanças de turno conversacional: *mas*, reconhecida pela tradição para o primeiro uso listado e amplamente utilizada na fala espontânea; e *só que*, não contemplada pela tradição gramatical (talvez, por seu caráter inovador) e também empregada na fala espontânea<sup>3</sup>.

Na literatura sobre a alternância entre *mas* e *só que*, há trabalhos como o de Silva (2017), intitulado *A perífrase conjuncional só que: gramaticalização e variação linguística*, realizado no âmbito do curso de Mestrado em Linguística da Universidade Federal de Minas Gerais; e como o de Silva e Silva (2014), intitulado *Funcionalidade e gramaticalização da construção só que na fala goiana*, fruto do projeto de mesmo nome, integrado ao Programa Institucional Voluntário de Iniciação Científica da Universidade Federal de Goiás.

---

3 Para maior compreensão do percurso de gramaticalização da perífrase conjuncional “só que”, ver os trabalhos de Sanderléia Roberta Longhin-Thomazi (LONGHIN, 2002; LONGHIN-THOMAZI, 2003). Segundo a autora, ainda que o sentido básico, invariante, de “só que” seja sempre de quebra de expectativa, existem 5 diferentes acepções de sentido que podem ser assumidas pela construção, a saber: (a) marcador de diferença; (b) marcador de surpresa/acontecimento inesperado; (c) marcador de refutação; (d) marcador de contra-argumentação; (e) marcador de não-ratificação de condições. Vale destacar que, em todas essas acepções, a perífrase conjuncional “só que” é intercambiável por “mas”.

Silva (2017) assume os objetivos de traçar o processo de gramaticalização da partícula *só que*; e de verificar sua implementação na variedade de Português Brasileiro falada em Belo Horizonte (MG), representada por amostras de fala informal do Projeto C-ORAL-BRASIL, compilado por pesquisadores da Universidade Federal de Minas Gerais. Por suas vezes, Silva e Silva (2014) têm o objetivo de investigar a alternância entre *mas* e *só que* na variedade de Português Brasileiro falada em Goiás, representada por amostras de fala do *corpus* reunido por pesquisadores do Grupo de Estudos Funcionalistas da Universidade Federal de Goiás.

O estudo de Silva (2017) tem dois momentos: inicialmente, busca analisar a trajetória de gramaticalização da perífrase conjuncional *só que*; e, feito isso, realiza um estudo variacionista para verificar a estabilidade dessa forma na variedade de Português Brasileiro falada em Belo Horizonte, ensejando sua alternância com a conjunção *mas*.

Silva (2017) selecionou um total de 36 entrevistas, totalizando 20 falantes homens e 20 falantes mulheres. O *corpus* do qual essas entrevistas foram recortadas está estratificado por faixa etária (A, de 18 a 25 anos; B, de 26 a 40 anos; C, de 41 a 59 anos; e D, acima de 60 anos); por gênero (masculino e feminino) e por nível de escolaridade (1, até sete anos; 2, graduados). Cinco grupos de fatores foram considerados pela pesquisa: três de caráter social, de acordo com os quais o *corpus* adotado foi estratificado (gênero, faixa etária, nível de escolaridade); e um de caráter linguístico, posição da perífrase conjuncional/conjunção (posição inicial de período; posição entre períodos). O processamento de dados, realizado com auxílio do pacote de programas Goldvarb X, demonstrou a existência de fatores que favorecem a aplicação da regra variável, o uso da perífrase conjuncional. Foram obtidos, respectivamente, os pesos relativos .66 para a faixa etária A e .63 para a faixa D; e .58 para a posição entre orações.

Por sua vez, o trabalho de Silva e Silva (2014) investiga a alternância entre *mas* e *só que* em amostras de fala da variedade de Português Brasileiro falada em Goiás. Foram analisadas 18 entrevistas, que estão estratificadas por gênero, por faixa etária (I, de 15 a 25 anos; II, de 26 a 36 anos; e III de 37 a 55 anos). Ao todo, são 9 homens e 9 mulheres. Não é explicitada a distribuição de homens e de mulheres pelos fatores em relação aos quais o *corpus* é controlado.

A análise de Silva e Silva (2014) conclui que a perífrase conjuncional *só que* compartilha com a conjunção *mas* os valores opositivo e de contraexpectativa, mas que, ao contrário dela, pode ter, também, a função de operador de foco, oriunda do item *só*. De uma perspectiva sintática, *só que* é avaliado pelos autores como item conjuncional, na medida em que coordene sentenças. De uma perspectiva pragmático-discursiva, consideram-no tanto como operador de foco, nos casos em que direcione a atenção do interlocutor para o conteúdo que a segue, quanto como operador discursivo, nos casos em que introduza argumentos discursivamente mais proeminentes do que os que já foram mencionados

no discurso (LONGHIN, 2002; LONGHIN-THOMAZI, 2004); essa segunda função é compartilhada com a conjunção *mas*.

## Pressupostos teóricos e metodológicos

Com o objetivo de apresentar um mapeamento sincrônico da variação entre as formas *mas* e *só que* nas variedades de Português Brasileiro faladas nas cidades de Nova Iguaçu e do Rio de Janeiro, mostra-se pertinente a adoção dos pressupostos teórico-metodológicos da Sociolinguística de base laboviana (LABOV, 1994, 2003). Partindo de tal perspectiva, os dados linguísticos em estudo foram levantados, com base em entrevistas orais coletadas entre 36 entrevistas sociolinguísticas, do tipo entrevistador-informante, previamente gravadas e transcritas.

Os dados linguísticos localizados em tais entrevistas foram codificados de acordo com grupos de fatores linguísticos e sociais, e, posteriormente, submetidos ao Programa estatístico de regras variáveis, denominado *Goldvarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005).

## Descrição dos *corpora* selecionados e números de ocorrências

Esta pesquisa se vale de duas amostras de fala do Corpus Concordância, disponibilizadas pelo Projeto CORPORAPORT<sup>4</sup>. Tais amostras foram coletadas no Brasil, nas regiões de Copacabana, bairro da cidade do Rio de Janeiro; e de Nova Iguaçu, cidade da baixada fluminense, região metropolitana do Estado do Rio de Janeiro. Apresentam-se na forma de 36 entrevistas sociolinguísticas, estratificadas em função da região geográfica, da faixa etária (de 18 a 35 anos; de 36 a 55 anos; e de 56 anos em diante), do nível de escolarização (Ensino Fundamental, Ensino Médio e Ensino Superior) e do gênero do informante (feminino e masculino).

Perceba-se que os dois conjuntos de amostras de fala são simétricos, os quatro grupos de fatores de acordo com os quais as amostras são estratificadas cruzam-se da seguinte maneira: por região geográfica, para cada faixa etária, há dois falantes de cada faixa de escolaridade, um de cada gênero, totalizando dezoito entrevistas no Rio de Janeiro e dezoito entrevistas em Nova Iguaçu.

---

4 O Projeto CORPORAPORT é a sequência do Projeto *Padrões de Concordância em variedades africanas, brasileiras e europeias do português*, responsável pela organização do *Corpus Concordância*, cujas subamostras de Nova Iguaçu e Copacabana são usadas na presente investigação.

## Hipóteses e grupos de fatores condicionantes investigados

De acordo com Silva (2017, p. 73), a perífrase conjuncional *só que* “não parece ser estigmatizada socialmente, na medida em que é usada livremente por homens e mulheres de diversas faixas etárias”. Partindo das observações da autora, aventam-se as seguintes hipóteses:

1. É notável que a perífrase conjuncional *só que* não seja estigmatizada. Isso pode ser um indício de que fatores sociais que normalmente são avaliados como negativos (marcas associadas a falantes de baixa escolaridade, por exemplo) não atuem na alternância entre *mas* e *só que*. Portanto, aproveitando o ensejo de que os *corpora* adotados são já estratificados por escolaridade, por faixa etária, por gênero e por região geográfica, é importante verificar se cada um desses fatores é ou não relevante para o fenômeno.

2. Silva (2017), talvez pela baixa incidência de ocorrências para o fenômeno, não foi capaz de encontrar a motivação para a alternância *mas* e *só que*. Acontece, porém, que a autora não considera fatores discursivos em sua análise. Portanto, é imprescindível tentar determinar a função discursiva de cada uma das formas — será que ocorrem, livremente, nos mesmos contextos? Será, ao contrário, que estão em distribuição complementar — ocorrem nos mesmos contextos, mas dependem de motivações diferentes? —, o que servirá para esclarecer se um tratamento variacionista para esses usos é o mais adequado; no caso de ser, servirá, também, para esclarecer qual é a motivação do uso de uma e de outra forma.

Para testar essas hipóteses, foi considerado um conjunto de oito grupos de fatores, entre sociais e linguísticos, a saber: região geográfica, gênero, faixa etária, nível de escolaridade, tipo de construção sintática, transitividade verbal (HALLIDAY, 1985), sequência tipológica e negação.

## Análise dos resultados

Nesta seção, são apresentados os resultados da análise quantitativa, percentuais e peso relativo, os quais são submetidos a uma interpretação qualitativa, com a finalidade de esclarecer (i) qual é o estatuto das formas variáveis nas variedades estudadas; e (ii) quais são os fatores que favorecem e que inibem o uso de cada uma das formas. Para todos os resultados, o fator de aplicação é a variante *só que*. Na designação da melhor rodada, o programa estatístico selecionou, dentre todos, os fatores *construção sintática* e *transitividade verbal* como sendo os de maior relevância. Quanto aos fatores não selecionados pelo programa, são oferecidos os resultados das rodadas individuais.

A ordem de apresentação dos fatores, a seguir, não é a ordem em que foram selecionados pelo programa; em vez disso, os fatores sociais vêm à frente dos fatores linguísticos, na intenção de fornecer, primeiramente, um panorama dos aspectos sociais do fenômeno. Em seguida, há a apresentação dos fatores linguísticos.

### **A faceta social do fenômeno variável: atuação apagada da região geográfica, do gênero e da escolaridade, em contraste com a aparente relevância da faixa etária**

Quando se tem em vista o controle da *região geográfica*, é possível verificar que, do total de ocorrências do Rio de Janeiro, 488 ocorrências, 44 (9,0%) são de *só que* e 444 (91,0%) são de *mas*. As 488 ocorrências representam 41% do número total de ocorrências (1169). Estritamente quanto ao percentual de ocorrências das variantes, *mas* é, com grande vantagem, a mais utilizada entre as duas. No que tange ao peso relativo da variedade do Rio de Janeiro sobre a implementação do fenômeno, verifica-se que pertencer a essa comunidade de fala não é um fator que impacta na opção por uma das duas variáveis: o valor do peso relativo é .50, o índice de neutralidade.

Para Nova Iguaçu, foram registradas 681 ocorrências de *mas* e de *só que*: 58 (8,5%) ocorrências de *só que* e 623 (91,5%) ocorrências de *mas*. As 681 ocorrências representam 59% do número total de ocorrências. Como para o Rio de Janeiro, com percentuais bastante próximos, *mas* é a variante preferida pelos falantes dessa variedade. Quanto ao peso relativo, o valor associado pela análise estatística, em sua melhor rodada, à variedade de Nova Iguaçu é, também, semelhante ao atribuído à variedade do Rio de Janeiro: .49. Portanto, pertencer à comunidade de fala de Nova Iguaçu não é, do mesmo modo, um fator que impacta na opção por uma das variáveis. Na Tabela 1, são trazidas as informações contidas neste e no parágrafo anterior (número de ocorrências, percentuais e pesos relativos).

**Tabela 1.** A influência da origem geográfica na aplicação de “só que”

| Rio de Janeiro |     |     | Nova Iguaçu |     |     |
|----------------|-----|-----|-------------|-----|-----|
| NO             | %   | PR  | NO          | %   | PR  |
| 44/488         | 9,0 | .50 | 58/681      | 8,5 | .49 |

**Fonte:** Elaboração própria

A partir da análise da influência da origem geográfica dos falantes, é possível afirmar que a circunscrição da alternância entre *mas* e *só que* não é a uma variedade específica de Português Brasileiro — pelo menos, não às variedades observadas nesta pesquisa; se for, o recorte das variedades é mais amplo e envolve grandeza como “Português falado no Estado do Rio de Janeiro” *versus* “Português Brasileiro falado no Estado da Bahia”,

por exemplo; isto é, por toda sua especificidade, uma hipótese que deverá ser verificada por pesquisas futuras, a partir de *corpora* que abarquem outras variedades de Português Brasileiro. Portanto, se a opção ou por *mas* ou por *só que* não é motivada por características sociolinguísticas exclusivas das variedades observadas (e isso é corroborado pelos pesos relativos traçados para cada uma delas, que beiram à neutralidade), isso pode estar relacionado ao fato de que a forma inovadora (*só que*) não é estigmatizada.

No que se refere ao controle do *gênero*, constam, do total de 1169 ocorrências de *mas* e de *só que*, 587 ocorrências para homens e 582 ocorrências para mulheres. Das 587 ocorrências do gênero masculino, 43 (7,3%) são de *só que* e 544 (92,7%) são de *mas*; e, das 582 ocorrências do gênero feminino, 59 (10,1%) são de *só que* e 523 (89,9%) são de *mas*. Se forem somadas, as ocorrências de *só que* de ambos os gêneros dão 102 e representam 8,7% das 1169 ocorrências, ao passo que as ocorrências de *mas* dão 1067 e representam 91,3% das 1169 ocorrências. Na aferição do peso relativo individual de ambos os gêneros para a aplicação da regra, nenhum é apontado pela análise estatística como altamente relevante: o peso relativo do gênero masculino é .45, e o do gênero feminino, .49. Confirmam essas informações, adiante, na Tabela 2.

**Tabela 2.** A influência do gênero na aplicação de “só que”

| Masculino |     |     | Feminino |     |     |
|-----------|-----|-----|----------|-----|-----|
| NO        | %   | PR  | NO       | %   | PR  |
| 43/587    | 7,3 | .45 | 59/582   | 8,5 | .49 |

**Fonte:** Elaboração própria

O fator nível de *escolaridade*, como se sabe, é um gatilho comum de fenômenos variáveis. Em função do fato de que indivíduos com diferentes graus de frequência aos canais de instrução formal podem vir a utilizar formas linguísticas consideradas de maior prestígio do que outras, devido à exposição maior às normas da variedade padrão de Português, formas linguísticas que são consideradas marcas dialetais ou de variedades específicas. Considerando que a alternância entre *mas* e *só que* não seja estigmatizada, isso pode ser um indício de que não seja motivada pelo fator escolaridade, de modo que falantes de diferentes níveis de acesso à instrução formal deverão ter percentuais de preferência pela aplicação da regra variável idênticos. Veja-se, a seguir, se essa expectativa foi confirmada.

A distribuição das 102 ocorrências de *só que* e das 1067 ocorrências de *mas* pelos níveis de escolaridade é a seguinte: para o nível de escolaridade 1, que engloba falantes cujos anos de frequência à instrução formal variam entre 6 e 9 (primeiro segmento do Ensino Fundamental completo; segundo segmento do Ensino Fundamental completo ou incompleto), há 41 (10,5%) ocorrências de *só que* e 348 (89,5%) ocorrências de *mas*; para o nível de escolaridade 2, que engloba falantes cujos anos de frequência à instrução formal

variam entre 10 e 12 (Ensino Médio completo ou incompleto), há 26 (7,7%) ocorrências de *só que* e 312 (92,3%) ocorrências de *mas*; e, para o nível de escolaridade 3, que engloba falantes cujos anos de frequência à instrução formal são superiores a 13 (Ensino Superior completo ou incompleto), há 35 (7,9%) ocorrências de *só que* e 407 (92,1%) ocorrências de *mas*. O peso relativo associado pela análise estatística, na rodagem individual, aos três níveis de escolaridade considerados não destaca nenhum deles como especialmente favorável à aplicação da regra variável: .55 é o peso relativo do nível de escolaridade 1, muito próximo ao patamar de neutralidade; .46 é o peso relativo do nível de escolaridade 2, também muito próximo ao patamar de neutralidade; e .47 é o peso relativo do nível de escolaridade 3, que, semelhantemente aos outros, é muito próximo ao patamar de neutralidade. Confirmam-se essas informações na Tabela 3, abaixo.

**Tabela 3.** A influência do nível de escolaridade na aplicação de “só que”

| Escolaridade 1<br>(Ensino Fundamental) |      |     | Escolaridade 2<br>Ensino Médio |     |     | Escolaridade 3<br>Ensino Superior |     |     |
|--|------|-----|--------------------------------|-----|-----|-----------------------------------|-----|-----|
| NO                                     | %    | PR  | NO                             | %   | PR  | NO                                | %   | PR  |
| 41/389                                 | 10,5 | .55 | 26/338                         | 7,7 | .46 | 35/442                            | 7,9 | .47 |

**Fonte:** Elaboração própria

Os resultados para *faixa etária*, por sua vez, indicam um total de 102 (8,7%) ocorrências para *só que* e de 1067 (91,3%) ocorrências para *mas*, distribuídas da seguinte maneira: para *só que*, há 35 (34,3%) ocorrências para faixa etária A, 47 (46,1%) ocorrências para a faixa etária B e 20 (19,6%) ocorrências para a faixa etária C; e, para *mas*, há 403 (37,8%) ocorrências para a faixa etária A, 326 (30,5%) ocorrências para a faixa etária B e 358 (31,7%) ocorrências para a faixa etária C. Os pesos relativos atribuídos pela rodagem individual às faixas etárias A, B, C são os seguintes: para a faixa etária A, .48; para a faixa etária B, .61; e, para a faixa etária C, .39.

A partir desses pesos relativos, verifica-se que pertencer à faixa etária A, a que alberga os falantes de menor idade, não é um fator que favorece a aplicação da regra variável; contudo, a comparação entre os pesos relativos associados às faixas etárias B e C demonstram que pertencer à faixa etária B, que alberga os falantes de idade intermediária, é um fator que favorece ligeiramente a aplicação de *só que*, ao passo que pertencer à faixa etária C, que alberga os falantes de maior idade, é um fator que desfavorece, também ligeiramente, a aplicação da regra variável. Tudo isso parece apontar na mesma direção para a qual a análise de Ferreira Silva (2017) sinaliza: a alternância entre *mas* e *só que* surgiu há cerca de quarenta e cinco anos, pouco antes da geração que corresponde à faixa C desta pesquisa, consolidando-se na fala da geração seguinte, que, por sua vez, tornou-se a geração B desta pesquisa. Confirmam-se essas informações na Tabela 4, que se segue.



**Tabela 4:** A influência da faixa etária na aplicação de “só que”

| Faixa etária A (18-35 ) |     |     | Faixa etária B (36-55) |      |     | Faixa etária C (56- ) |     |     |
|-------------------------|-----|-----|------------------------|------|-----|-----------------------|-----|-----|
| NO                      | %   | PR  | NO                     | %    | PR  | NO                    | %   | PR  |
| 35/438                  | 8,7 | .48 | 47/373                 | 14,4 | .61 | 20/378                | 5,6 | .39 |

**Fonte:** Elaboração própria

Se a alternância entre *mas* e *só que* pode ter surgido na geração que corresponde à faixa etária C desta pesquisa, o que explicaria o fato de ser justamente essa a geração que desfavorece a aplicação da regra variável? Uma explicação possível para isso é o fato de que essa alternância foi, em um primeiro momento, controlada pela geração em cuja fala surgiu; sendo assim, aventa-se que, de início, o fenômeno pode ter sido estigmatizado ou pode ter se tornado, à época, sinônimo de algum estrato da comunidade de fala, face ao que a comunidade de fala de então demonstrou uma postura conservadora. Como resultado, a atual faixa etária C exibe, inevitavelmente, as marcas linguísticas do período em que sua gramática se estabilizou, mas sua produção não favorece totalmente a aplicação da regra variável, característica que também reflete o comportamento conservador de parcela da comunidade de fala frente à variante *só que* nessa época.

Tendo em vista os resultados dos parâmetros sociais que estratificam a amostra, é possível depreender, de maneira geral, sua pouca influência em relação ao fenômeno variável, haja vista que nenhum deles tenha sido selecionado pelo Programa GoldVarb X. Todavia, há diferenças entre suas respectivas atuações: enquanto o controle da região geográfica, do gênero e da escolaridade se mostraram irrelevantes na escolha de uma ou outra forma linguística; a faixa etária parece ter alguma influência no fenômeno linguístico, em especial, o grupo etário intermediário (de 36 a 55 anos) parece impulsionar a escolha de “só que”.

Por fim, no que se refere à hipótese 1, vale destacar que a análise dos grupos de fatores sociais indica não se tratar de um processo de mudança em curso, e sim um caso de variação estável entre as formas linguísticas.

### **Os aspectos linguísticos do fenômeno variável: a influência determinante do tipo de construção sintática e da transitividade verbal**

O *tipo de construção sintática* é o fator de maior relevância selecionado pela análise estatística (à frente de transitividade verbal). Por detrás da seleção deste grupo de fatores, há a preocupação de se definir se a alternância entre *mas* e *só que* é condicionada, parcial ou totalmente, pelo tipo de contexto sintático em que ocorrem. A seguir, exemplos de *mas* e de *só que* para cada um dos tipos de construção sintática.

**Quadro 1.** Exemplos de tipos de construção sintática

| Tipo de construção sintática        | Exemplos  |
|-------------------------------------|---|
| conexão de sintagmas não oracionais | 1. não dentro da cidade... <b>mas nas vias</b> (NIG-A-2-H)<br>2. eu continuo na mesma empresa <b>só que em sistemas diferentes</b> (COP-B-3-H)  |
| conexão de sintagmas oracionais     | 3. eu converso com essa minha empregada... <b>mas chega uma hora que eu também quero ficar sozinha...</b> (COP-C-2-M)<br>4. não ganham nada em troca disso... <b>só que eles brigam</b> (COP-B-3-H)   |
| operador discursivo                 | 5. vieram falando que era polícia <b>mas até que... os ladrões: ...entre aspas até foram bastante POLIDOS</b> (NIG-C-3-M)<br>6. eu optei pela Estácio de Sá... que ela tinha comprado uma uma faculdade de arquitetura que iria abrir aí eu... me encaixei lá falei vou fazer administração... no decorrer do/ dos anos aí eu faço pra arquitetura... <b>só que ela nunca abriu né...</b> (COP-B-3-H) |
| marcador de turno conversacional    | 7. D: não mas só a gente vai ouvir a gente vai e depois pra tomar nota... ninguém vai ter acesso a isso<br>L: <b>mas pode ouvir não tem problema não</b> (NIG-C-3-H)<br>8. D: eles vendem então né<br>L: <b>só que lá no lixão...</b> (COP-C-1-H)   |

**Fonte:** Elaboração própria

A distribuição das 102 ocorrências de *só que* e das 1067 ocorrências de *mas* é a seguinte: para a função de relacionar sintagmas não oracionais, há 1 ocorrência de *só que* (4,2%) e 23 (95,8%) ocorrências de *mas*; para a função de relacionar sintagmas oracionais, há 65 (8,4%) ocorrências de *só que* e 709 (91,6%) ocorrências de *mas*; para a função de operador discursivo, definindo cadeias subtópicas no discurso, há 33 ocorrências de *só que* (9,7%) e 308 (90,3%) ocorrências de *mas*; e, para a função, de marcador de turno conversacional, há 3 (10%) ocorrências de *só que* e 27 (90%) ocorrências de *mas*. O peso relativo associado pela análise estatística, em sua melhor rodada, aos tipos de construções sintáticas nas quais *só que* e *mas* podem ocorrer selecionam dois fatores como relevantes para a aplicação da regra variável, os fatores marcador de turno (.73) e operador discursivo (.86); um fator como desfavorável à aplicação da regra, o fator relação de sintagmas não oracionais (.11); um fator como neutro à aplicação da regra, o fator sintagma oracional (.51). Confirmam-se essas informações, abaixo, na Tabela 5.

**Tabela 5.** A influência do tipo de construção sintática na aplicação de “só que”

| Sintagmas não oracionais |     |     | Sintagmas oracionais |     |     | Operador Discursivo |     |     | Marcador de turno |    |     |
|--------------------------|-----|-----|----------------------|-----|-----|---------------------|-----|-----|-------------------|----|-----|
| NO                       | %   | PR  | NO                   | %   | PR  | NO                  | %   | PR  | NO                | %  | PR  |
| 1/24                     | 4,2 | .11 | 65/774               | 8,4 | .51 | 33/442              | 9,7 | .86 | 3/30              | 10 | .73 |

Fonte: Elaboração própria

Com relação à hipótese 2, quais são os condicionamentos linguísticos que favorecem a aplicação de *só que*, nota-se que as funções de operador discursivo e de marcador de turno conversacional, que têm proeminência no discurso, são as que mantêm estreita relação com *só que*. Portanto, parece que se delineie um quadro em que os falantes utilizam a forma *mas* para veicular informações de pouca relevância no contexto discursivo (contextos habituais de adversidade), enquanto preferem a forma *só que* nos contextos em que a relevância da informação é maior (contextos em que há construção de cadeias subtópicas, indicando que a informação seguinte ou é nova ou é importante; contextos em que o falante toma para si o turno de fala, inserindo uma nova informação no discurso). Nesse sentido, tal resultado referenda o que trabalhos anteriores propõem em relação às especificidades semântico-pragmáticas que caracterizam a forma inovadora. Nas palavras de Longhin (2002, p. 142),

*Só que* funciona como um operador de foco, que acrescenta uma circunstância em geral nova, não considerada até o momento, e que apresenta esta circunstância como sendo a única que se justifica adicionar. Esse elemento novo contrasta com tudo mais no tipo de conclusão que autoriza e é suficiente para tomar inválida uma generalização previamente considerada.

Há, abaixo, três exemplos de *só que* (9-11) e três exemplos de *mas* (12-14) em que a proeminência da informação veiculada no discurso está em jogo.

9. ele falou “oh vocês vão ter um curso básico de iluminação” **só que eu já tinha feito uma iluminação...** (NIG-A-1-H)
10. L: assim matéria... gostava mais de português... e espanhol D: espanhol é legal... L: **só que eu não aprendi nada** (COP-A-1-M)
11. ela tava aparentemente bem **só que... ela bateu com a cabeça...** (COP-A-2-M)

12. eu sou uma idosa **mas eu trabalho** (COP-C-1-M)
13. o colégio era pago **mas tive tantos problemas** (NIG-B-M)
14. o Euclides da Cunha ele escreveu um livro chamado os sertões... **mas é chato para cacete...** (NIG-C-3-H)

O segundo grupo de fatores selecionado pelo programa estatístico diz respeito ao controle da *transitividade verbal*. A partir de tal variável, na acepção de Halliday (1985), controla-se se o tipo de atividade representada pelos verbos — é relevante para a alternância entre *mas* e *só que* (por exemplo, é possível verificar se o tipo X de verbo propicia o uso da forma Y).

O tratamento de Halliday (1985) para a transitividade verbal supõe que as diferentes classes semânticas nas quais se dividem os verbos funcionam como uma representação da experiência dos indivíduos; esse foi o segundo e último grupo de fatores selecionado pela análise estatística. A partir dessa perspectiva, transitividade verbal não é entendida em termos de complementação verbal, mas é vista como a materialização, por meio da semântica associada aos verbos, da experiência acumulada pelos falantes. Por esta análise, foram consideradas todas as seis categorias de transitividade verbal propostas por Halliday (processos materiais, processos mentais, processos relacionais, processos verbais, processos existenciais e processos comportamentais) e, além dessas, a possibilidade de que a construção sintática em que a perífrase conjuncional/conjunção ocorre não conter verbo, por não ser oracional. Abaixo, exemplos de *mas* e de *só que* em função dos diferentes tipos de transitividade verbal.

**Quadro 2.** Exemplos de transitividade verbal

| Tipo de verbo     | Característica   | Exemplo   |
|-------------------|--|---|
| processo material | Designam ações e processos que desencadeiam mudanças externas. | 15. eles botaram na teflon só que eu <b>botei esse papel</b> (COP-C-3-M)<br>16. ele: ... faz as melhorias <b>mas também... ele retira do povo ...</b> (NIG-A-3-H)   |
| processo mental   | Designam processos cognitivos e de percepção do mundo.         | 17. Posso ter meus problemas <b>só que eu penso assim... meus problemas eu tenho que resolver dentro da minha casa</b> (NIG-B-1-M)<br>18. é uma equipe que anda fazendo isso <b>mas eu não sabia...</b> (COP-C-1-M) |

|                           |   |   |
|---------------------------|---|---|
| processo relacional       | Designam relações entre entidades.  | 19. ganhamos outras causas <b>só que a essa altura do campeonato eu <u>era</u> muito amiga de um dos engenheiros do Banco do Brasil...</b> (COP-C-3-M)<br>20. eu nem conheço esse hospital <b>mas a cirurgia dele é muito boa</b> (NIG-B-1-M)   |
| processo verbal           | Designam atividades do comunicar.   | 21. existe uma forma certa de falar... <b>mas geralmente você nunca vai <u>falar</u> exatamente o que você escreve...</b> (COP-A-2-M)<br>22. as pessoas me procuram pra mim escrever alguma coisa... <b>só que eu <u>falo</u> pra eles "oh...eu vou escrever e vou passar pelo professor de português... senão já era"...</b> (NIG-B-1-H) |
| processos existenciais    | Designam a existência ou o acontecimento de algo; contam com um único participante. | 23. não tenho nada para falar do ensino público adoro... <b>só que à noite <u>tem</u> muita gente que vai para brincar...</b> (NIG-A-2-M)<br>24. houve isso <b>mas no meu ano graças a Deus não <u>houve</u> incidentes maiores...</b> (COP-C-3-H)  |
| processos comportamentais | Designam comportamentos e atividades fisiológicas e psicológicas.                   | 25. a menina que tava na minha frente ela tá com (os) mesmo(s) sintomas que eu tou... <b>só que ela também/ ela <u>começou a evacuar sangue</u></b> (NIG-A-2-M)<br>26. dependendo da forma que você... tirar raio-x na pessoa você pega um câncer <b>mas eu nunca <u>ouvi</u> falar nisso na minha vida né...</b> (NIG-A-2-M)             |
| sem verbo                 |   | 27. uma repulsa não pelos indivíduos <b>mas pela condição</b> (COP-A-2-H)<br>28. eu continuo na mesma empresa <b>só que em sistemas diferentes</b> (COP-B-3-H)  |

Fonte: Elaboração própria

A distribuição das 102 ocorrências de *só que* e das 1067 ocorrências de *mas* é a seguinte: para verbos que denotam processos materiais, há 29 (9,3%) ocorrências de *só que* e 284 (90,7%) ocorrências de *mas*; para verbos que denotam processos mentais, há 15 (10,9%) ocorrências de *só que* e 122 (89,1%) ocorrências de *mas*; para verbos que denotam processos relacionais, há 26 (6,6%) ocorrências de *só que* e 366 (93,4%) ocorrências de *mas*; para processos verbais, há 8 (18,2%) ocorrências de *só que* e 36 (81,8%) ocorrências de *mas*; para processos existenciais, há 9 (8,3%) ocorrências de *só que* e 99 (91,7%) ocorrências de *mas*; para processos comportamentais, há 12 (9%) ocorrências de *só que* e 122 (91%) ocorrências de *mas*; e, para construções sintáticas não oracionais, em que não há verbo, há 3 (7,3%) ocorrências de *só que* e 38 (92,7%) de ocorrências de *mas*.

O peso relativo associado pela análise estatística, em sua melhor rodada, aos tipos de transitividade verbal selecionam três fatores como relevantes para a aplicação da regra

variável, verbos que denotam processos mentais (.71), verbos que denotam processos materiais (.74) e verbos que denotam processos verbais (.88); um fator como desfavorável à aplicação da regra, construções sintáticas não oracionais (.22); e três fatores neutros à aplicação da regra, verbos que denotam processos relacionais (.43), verbos que denotam processos existenciais (.49) e verbos que denotam processos comportamentais (.51) Confirmam-se essas informações, abaixo, na Tabela 6.

**Tabela 6.** A influência da transitividade verbal na aplicação de “só que”

| Processos verbais      |      |     | Processos materiais   |     |     | Processos mentais |      |     | Processos comportamentais |   |     |
|------------------------|------|-----|-----------------------|-----|-----|-------------------|------|-----|---------------------------|---|-----|
| NO                     | %    | PR  | NO                    | %   | PR  | NO                | %    | PR  | NO                        | % | PR  |
| 8/30                   | 18,2 | .88 | 29/333                | 9,3 | .74 | 15/137            | 10,9 | .71 | 12/134                    | 9 | .51 |
| Processos existenciais |      |     | Processos relacionais |     |     | sem verbo         |      |     |                           |   |     |
| NO                     | %    | PR  | NO                    | %   | PR  | NO                | %    | PR  |                           |   |     |
| 9/108                  | 8,3  | .49 | 26/392                | 6,6 | .43 | 3/41              | 7,3  | .22 |                           |   |     |

Fonte: Elaboração própria

No que se refere à hipótese 2 (Quais são os condicionamentos linguísticos que favorecem a aplicação de *só que?*), o que parece haver em comum entre os tipos de transitividade verbal selecionados como favoráveis à aplicação de *só que* é o fato de que, nos processos denotados por ele, é comum haver mais de uma entidade envolvida, ao contrário dos demais tipos de transitividade verbal. Veja-se, nos exemplos acima, que verbos de processos verbais envolvem uma entidade que produz algum tipo de discurso e um discurso que é produzido; que verbos de processos materiais ensejam tanto uma entidade que é agente quanto uma entidade sobre a qual recai uma ação; e que verbos de processos mentais ensejam tanto uma entidade que processa uma informação quanto uma entidade que é processada. Portanto, a conclusão que parece emergir disso é que esses tipos de transitividade verbal permitem o encaixamento de entidades (isto é, de novas informações) na estrutura sintática de um determinado tipo verbal, e é em direção à especificação de *só que* por esse tipo de contexto que esta análise segue, também, quanto a outros fatores.

### **Outros grupos de fatores linguísticos controlados: sequência tipológica e contexto de negação.**

Cada instância de uso das línguas humana está atrelada a (i) um gênero discursivo e a (ii) uma tipologia textual. O gênero discursivo que é tomado por esta pesquisa para ser investigado decorre dos *corpora* selecionados; contudo, as *sequências tipológicas* dos enunciados dependem, exclusivamente, da decisão dos falantes. Desse modo, as sequências tipológicas indicam tendências a respeito do comportamento dos falantes e se tornam um fator relevante de análise. A partir dessa variável, controla-se se a

sequência tipológica condiciona o uso de algumas das formas alternantes. As sequências tipológicas estão dispostas, com exemplos, no Quadro 3, adiante.

**Quadro 3.** Exemplos de sequências tipológicas

| Sequência tipológica | Exemplos   |
|----------------------|--|
| narração             | 29. até aquele momento eu não tinha visto nada que fosse suspeito <b>mas tinha um cara do meu lado ele disse que seríamos assaltados</b> (COP-A-2-M)<br>30. aí minha mãe não pôde ir comigo que ela tava na fisioterapia foi a minha tia... <b>só que eu tava passando muito mal</b> (NIG-A-2-M)   |
| descrição            | 31. minha filha entra chorando <b>mas elas brinca bate palma canta musiquinha bota música dança faz eles pular</b> (COP-A-1-M)<br>32. um jantar com as mulheres (do meu pai) <b>só que as mulheres (do meu pai) ficavam na ponta da mesa e eu (na outra ponta)</b> (COP-C-3-H)   |
| exposição            | 33. meu marido não ganha muito tá <b>só que eu trabalho para poder ajudar</b> (NIG-B-1-M)<br>34. eu adoro um cinema... adoro um teatro... <b>mas ele não gosta</b> (COP-C-2-M)   |
| argumentação         | 35. a cirurgia dele era muito boa <b>só que: de lá para cá caiu muito</b> (NIG-B-1-M)<br>36. Lula: é é uma pessoa muito emblemática ele tem um aspecto camaleão que a gente tem que reconhecer <b>mas o Hitler também tinha... tá?</b> (COP-B-3-M)   |
| injunção             | 37. Ele disse assim “envia amor”... eu disse “envia amor? o quê que é isso?”... fui pra casa sentei comecei a escrever... o quê que eu escrevia?... era o Espírito Santo escrevendo por mim né... o quê que eu escrevia? Ele dizia assim... “fulano... você é perfeito porque Jesus/ Deus fez você à semelhança dele... você é bom honrado direito... você é fantástico você é precioso pra Deus”... então eu só mandei coisas bonitas pra ele... nunca mais eu disse... “você bebe você não sei o que” não não não não não não... só “você é a semelhança de Deus foi assim que Deus fez você... você é a semelhança de Deus você tem todos/ tudo que Deus tem honrado de tudo você é... você que não sabe então você age mal”... então... <b>só que você não manda a carta...</b> eu peço ao Espírito Santo... que leve aquilo e coloque no coração da pessoa... (COP-C-1-M)<br>38. ganhe pouco... <b>mas faça o que você gosta...</b> (COP-C-1-M) |

**Fonte:** Elaboração própria

Os resultados para sequência apresentam a seguinte distribuição: para sequência tipológica descritiva, há 8 (11,8%) ocorrências de *só que* e 60 (88,2%) ocorrências de *mas*; para sequência tipológica argumentativa, há 51 (7,4%) ocorrências de *só que* e 636 (92,6%) ocorrências de *mas*; para sequência tipológica expositiva, há 21 (7,2%) ocorrências de *só que* e 272 (92,8) ocorrências de *mas*; para sequência tipológica narrativa, há 21 (17,8%)

ocorrências de *só que* e 97 (82,2%) ocorrências de *mas*; e, para sequência tipológica injuntiva, há 1 ocorrência de *só que* e 2 ocorrências de *mas*. Os pesos relativos atribuídos pela análise estatística, na rodada individual, às cinco sequências tipológicas são os seguintes: favorecendo a aplicação da regra variável, há descrição, com .91 e injunção, com .94 e, neutras à aplicação da regra variável, há descrição, com .59, argumentação, com .46, e exposição, com .45. Faz-se, aqui, uma ressalva ao alto peso relativo associado à sequência tipológica injuntiva, visto que sua frequência nos *corpora*, dentre as 1169 ocorrências analisadas, resume-se a três dados. Confirmam-se essas informações na Tabela 7, a seguir.

**Tabela 7.** A influência da sequência tipológica na aplicação de “só que”

| injunção     |      |     | narração  |      |     | descrição |      |     |
|--------------|------|-----|-----------|------|-----|-----------|------|-----|
| NO           | %    | PR  | NO        | %    | PR  | NO        | %    | PR  |
| 1/3          | 33,3 | .94 | 21/118    | 17,8 | .88 | 8/68      | 11,8 | .59 |
| argumentação |      |     | exposição |      |     |           |      |     |
| NO           | %    | PR  | NO        | %    | PR  |           |      |     |
| 51/687       | 7,4  | .46 | 21/293    | 7,2  | .45 |           |      |     |

Fonte: Elaboração própria

Por fim, o último grupo de fatores controlado na amostra refere-se aos *contextos de negação*. Antes de prosseguir aos resultados, relembrem-se os exemplos de *mas* e de *só que* em função dos contextos de negação.

**Quadro 4.** Exemplos de negação

| Contexto negativo  | Exemplos   |
|--|--|
| partícula de negação no segmento anterior à conjunção/perífrase conjuncional       | 39. eu <b>não</b> gosto de beber... às vezes eu vou <b>mas</b> ele gosta (COP-C-2-M)<br>40. <b>não</b> ganham nada em troca disso... <b>só que</b> eles brigam (COP-B-3-H)                         |
| partícula de negação no segmento introduzido pela conjunção/perífrase conjuncional | 41. eu sou mais velha que vocês <b>mas</b> eu não sou tão velha (NIG-B-1-M)<br>42. esquece que a criança tem que comprar livro e vai botando <b>só que</b> aí <b>não</b> tem condições (COP-C-1-H) |



|   |   |
|---|---|
| partícula de negação tanto no segmento anterior à conjunção/perífrase conjuncional quanto no segmento introduzido pela conjunção/perífrase conjuncional | 43. <b>não</b> sei o que que foi... <b>mas</b> eu <b>nunca</b> consegui seguir muitas regras <b>não</b> (NIG-A-2-H)<br>44. no caso uma parte vai e a outra <b>não só que</b> normalmente também <b>não</b> tem restrições (COP-B-3-H) |
| ausência de negação   | 45. Nova Iguaçu é uma cidade... como qualquer uma outra do Rio de Janeiro... só que: tem seus pontos positivos e seus pontos negativos... (NIG-A-2-H)<br>46. eu gosto <b>mas sacaneio eles o dia todo...</b> (COP-C-2-H)              |

Fonte: Elaboração própria

Os resultados para o contexto de negação têm a distribuição seguinte: para construções em que não há negação, há 72 (10,6%) ocorrências de *só que* e 607 (89,4%) ocorrências de *mas*; para construções em que há negação no segmento anterior à conjunção/perífrase conjuncional, há 7 (3,5%) ocorrências de *só que* e 193 (96,5%) ocorrências de *mas*; para construções em que há negação no segmento introduzido pela conjunção/perífrase conjuncional, há 22 (8,3%) ocorrências de *só que* e 244 (91,7%) ocorrências de *mas*; e, para construções em que há negação tanto no segmento anterior quanto no segmento introduzido pela conjunção/perífrase conjuncional, há 1 (4,2%) ocorrência de *só que* e 23 (95,8%) ocorrências de *mas*. A análise estatística não selecionou nenhum fator desse grupo como relevante à aplicação da regra variável; dois fatores, negação em segmento anterior e negação nos segmentos anterior e posterior, foram apontados como desfavoráveis à aplicação da regra variável. Os valores obtidos para peso relativo, na rodada individual, são os seguintes: ausência de negação, .57; negação no segmento anterior à conjunção/perífrase conjuncional, .28; negação no segmento introduzido pela conjunção/perífrase conjuncional, .50; negação nos segmentos anterior e posterior conjunção/perífrase conjuncional, .32. Confirmam-se essas informações na Tabela 8.

**Tabela 8.** A influência do contexto de negação na aplicação de “só que”

| ausente |      |     | segmento anterior |     |     | segmento posterior |     |     | ambos os segmentos |     |     |
|---------|------|-----|-------------------|-----|-----|--------------------|-----|-----|--------------------|-----|-----|
| NO      | %    | PR  | NO                | %   | PR  | NO                 | %   | PR  | NO                 | %   | PR  |
| 72/679  | 10,6 | .57 | 7/200             | 3,5 | .28 | 22/266             | 8,3 | .50 | 1/224              | 4,2 | .32 |

Fonte: Elaboração própria

## Considerações finais

Dos oito grupos de fatores linguísticos considerados — origem geográfica, gênero, nível de escolaridade, faixa etária, tipo de construção sintática, transitividade verbal,

sequência tipológica e negação —, é importante destacar, de início, que apenas dois (tipo de construção sintática e transitividade verbal) foram, de fato, selecionados como relevantes pelo programa estatístico de regras variáveis, o *Goldvarb X*.

Em relação aos contextos linguísticos e discursivos que favorecem a ocorrência de *só que*, esses são: construções sintáticas em que *só que* funciona ou como operador discursivo ou como marcador de turno conversacional; coadunado com verbos que denotam processos materiais, processos verbais e processos mentais; e em sequências tipológicas narrativas (e injuntivas). Sendo assim, satisfazendo a hipótese 2, conclui-se, a partir disso, que *só que* parece ser especializado em veicular informações que os falantes desejam destacar de alguma maneira ou informações que ainda não tenham sido introduzidas no discurso, enquanto *mas* é empregado nos demais contextos. Tal resultado referenda a análise de Longhin-Thomazi (2004, p. 237) no que se refere à construção inovadora: “*só que* se comporta como um legítimo marcador de foco, entendendo-se por ‘foco’, nos critérios da Gramática Funcional, a função pragmática que quando indicada a um constituinte confere a ele o estatuto de informação mais importante”.

## REFERÊNCIAS

HALLIDAY, M. A. K. *An Introduction to Functional Grammar*. Londres: Edward Arnold, 1985.

LABOV, W. Some Sociolinguistic Principles. In: PAULSTON, C. B.; TUCKER, G. R. (ed.). *Sociolinguistics: The Essential Readings*. Massachusetts: Blackwell, 2003.

LABOV, W. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell, 1994.

LONGHIN, S. R. *A gramaticalização da perífrase conjuncional ‘só que’*. 2002. Tese (Doutorado em Língua Portuguesa) – Instituto de Estudos de Linguagem, Universidade Estadual de Campinas, Campinas, 2002.

LONGHIN-THOMAZI, S. R. A gramaticalização da perífrase conjuncional “só que”. *Estudos Linguísticos*, v. 33, p. 232-237, 2004.

LONGHIN-THOMAZI, S. R. A perífrase conjuncional “só que”: invariância e variantes. *ALFA: Revista de Linguística*, v. 47, n. 2, p. 139-152, 2003.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SILVA, P. H.; SILVA, L. A. *Funcionalidade e gramaticalização da construção “só que” na fala goiana*. Relatório final do Programa Instituição Voluntário de Iniciação Científica da Universidade Federal de Goiás, Goiânia, 2014.

SILVA, L. F. F. *A perífrase conjuncional “só que”*: gramaticalização e variação linguística. 2017. Dissertação (Mestrado em Linguística teórica e descritiva) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2017.

VIEIRA, S. R.; MOTA, M. A. C. (org.). *Corpus Concordância*. Rio de Janeiro: Faculdade de Letras/UFRJ. In: VIEIRA, S. R.; BRANDÃO; S. F. *CORPORAPORT: Variedades do Português em análise*. Rio de Janeiro: Faculdade de Letras-UFRJ. Disponível em: [www.corporaport.lettras.ufrj.br](http://www.corporaport.lettras.ufrj.br). Acesso em: 18 abr. 2017.